②

# NAVAL POSTGRADUATE SCHOOL
## Monterey, California

AD-A204 365

DTIC
ELECTE
22 FEB 1989
S D
E

# THESIS

A BAYESIAN METHOD TO IMPROVE SAMPLING
IN WEAPONS TESTING

by

Theodore C. Floropoulos

December 1988

Thesis Advisor: Glenn F. Lindsay

Approved for public release; distribution is unlimited

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

# REPORT DOCUMENTATION PAGE

| | |
|---|---|
| 1a REPORT SECURITY CLASSIFICATION Unclassified | 1b RESTRICTIVE MARKINGS |
| 2a SECURITY CLASSIFICATION AUTHORITY | 3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited |
| 2b DECLASSIFICATION/DOWNGRADING SCHEDULE | |
| 4 PERFORMING ORGANIZATION REPORT NUMBER(S) | 5 MONITORING ORGANIZATION REPORT NUMBER(S) |
| 6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School — 6b OFFICE SYMBOL (If applicable) 55 | 7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School |
| 6c ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000 | 7b ADDRESS (City, State, and ZIP Code) Monterey, California 93943-5000 |
| 8a NAME OF FUNDING/SPONSORING ORGANIZATION — 8b OFFICE SYMBOL (If applicable) | 9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER |
| 8c ADDRESS (City, State, and ZIP Code) | 10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO |

11 TITLE (Include Security Classification)
A BAYESIAN METHOD TO IMPROVE SAMPLING IN WEAPONS TESTING

12 PERSONAL AUTHOR(S)
FLOROPOULOS, Theodore C.

| 13a TYPE OF REPORT | 13b TIME COVERED | 14 DATE OF REPORT (Year, Month, Day) | 15 PAGE COUNT |
|---|---|---|---|
| Master's Thesis | FROM ____ TO ____ | 1988 December | 53 |

16 SUPPLEMENTARY NOTATION The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| 17 COSATI CODES | | | 18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number) |
|---|---|---|---|
| FIELD | GROUP | SUB-GROUP | Prior Uniform Distribution, Bayesian Method, Confidence Interval Size |

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This thesis describes a Bayesian method to determine the number of samples needed to estimate a proportion or probabiity with 95% confidence when prior bounds are placed on that proportion. It uses the Uniform [a,b] distribution as the prior, and develops a computer program and tables to find the sample size. Tables and examples are also given to compare these results with other approaches to finding sample size. The improvement that can be obtained with this method's fewer samples, and consequently less cost in Weapons Testing is required to meet a desired confidence size for a proportion or probability.

| | |
|---|---|
| 20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS | 21 ABSTRACT SECURITY CLASSIFICATION Unclassified |
| 22a NAME OF RESPONSIBLE INDIVIDUAL Glenn F. Lindsay | 22b TELEPHONE (Include Area Code) 408-646-2688 — 22c OFFICE SYMBOL 55Ls |

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

* U.S. Government Printing Office 1986—606-24.



Unclassified

Approved for public release; distribution is unlimited.

# A Bayesian Method To Improve Sampling In Weapons Testing

by

Theodore C. Floropoulos
Lieutenant Commander, Hellenic Navy
B.S., Hellenic Naval Academy, 1972

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

*NAVAL POSTGRADUATE SCHOOL*
December 1988

Author: ______________________
Theodore Floropoulos

Approved by: ______________________
G. F. Lindsay, Thesis Advisor

______________________
Laura D. Johnson, Second Reader

______________________
Peter Purdue, Chairman,
Department of Operations Research

______________________
Kneale T. Marshall,
Dean of Information and Policy Sciences

# ABSTRACT

This thesis describes a Bayesian method to determine the number of samples needed to estimate a proportion or probability with 95% confidence when prior bounds are placed on that proportion. It uses the Uniform [a,b] distribution as the prior, and develops a computer program and tables to find the sample size. Tables and examples are also given to compare these results with other approaches for finding sample size. The improvement that can be obtained with this method is fewer samples, and consequently less cost in Weapons Testing is required to meet a desired confidence size for a proportion or probability.

TABLE OF CONTENTS

I. INTRODUCTION .................................... 1

II. SAMPLE SIZE TO ESTIMATE A PROPORTION USING THE CLASSICAL METHOD .................................... 4

A. THE POINT ESTIMATE FOR A PROPORTION .................................... 4

B. THE CONFIDENCE INTERVAL FOR A PROPORTION .................................... 6

C. DETERMINING THE SAMPLE SIZE FROM CONFIDENCE INTERVAL .................................... 7

III. THE BAYESIAN METHOD AND ESTIMATORS .................................... 12

A. BAYES' THEOREM .................................... 12

B. SELECTION OF THE PRIOR DISTRIBUTION .................................... 14

C. DERIVATION OF THE POSTERIOR .................................... 16

D. THE POSTERIOR DISTRIBUTION AND BAYES' THEOREM .................................... 19

IV. SAMPLE SIZE TO ESTIMATE A PROPORTION USING THE BAYESIAN METHOD .................................... 24

A. THE BAYESIAN CONFIDENCE INTERVAL .................................... 24

B. DETERMINING THE SAMPLE SIZE FROM THE BAYESIAN INTERVAL .................................... 26

C. TABLES FOR FINDING THE SAMPLE SIZE .................................... 29

V. SUMMARY AND SUGGESTIONS FOR FURTHER STUDY .................................... 39

A. SUMMARY .................................... 39

B. SUGGESTIONS FOR FURTHER STUDY .................................... 40

APPENDIX A THE APL PROGRAM "SAMPLE" USED TO COMPUTE SAMPLE SIZES FOR BAYESIAN INTERVALS WITH 95% CONFIDENCE LEVEL .................................42
APPENDIX B THE APL PROGRAMS USED TO COMPUTE THE INVERSE CDF OF A BETA DISTRIBUTED RANDOM VARIABLE .........43
LIST OF REFERENCES........................................45
INITIAL DISTRIBUTION LIST ..............................46

# ACKNOWLEDGEMENT

I would like to thank both Professor Glenn F. Lindsay and Terri LeCroix Turner. Their support, help, and encouragement with this thesis was invaluable.

# I. INTRODUCTION

"Probability is relative, in part to ... ignorance, in part to knowledge." [Ref. 1:p. 140]

This is the epitome of Laplace's interpretation of probability, stated in the 1951 translation of his book A Philosophical Essay On Probabilities. The topic of this thesis is to estimate a probability. In particular, we will try to answer the question of how many trials are necessary or what should be the sample size to estimate a proportion or probability from a set of Bernoulli trials.

In many forms of Weapon System testing, sampling is not done sequentially, and the number of items to be tested must be specified before testing begins. Clearly, enough weapon systems or components must be tested to furnish reasonable confidence in the resulting estimate of, say, system reliability. On the other hand, since testing is expensive and often destructive (e.g., missile launches), the sample size should be no larger than necessary.

Many measures of effectiveness for military systems are in the form of proportions, or probabilities of an attribute occurring. Some examples are

1. System Reliability,
2. Hit Probability,
3. Launch Probability,
4. Detection Probability, and
5. Fraction Defective.

In such cases, testing may often be described as performing a set of independent Bernoulli trials.

The problem is stated as follows: how many Bernoulli trials must we conduct, so that with a certain level of confidence, we can estimate the appropriate proportion or probability. A way to approach the problem is given by the definition of a confidence interval:

> A confidence interval for an unknown parameter gives an indication of the numerical value of our unknown parameter as well as a measure of how confident we are of that numerical value [Ref. 2:p. 323].

Given a desired confidence interval size for a proportion or probability, we wish to know the number of samples needed to provide a confidence interval of that size, and in this thesis we will produce tables and a computer program to assist in finding that sample size.

We will discuss two methods for the above calculations and we will compare the results. The first and well known one from classical statistics bases the estimate upon a simple random sample, and confidence intervals and sample size are explained in the next chapter. The second method, and primary focus of our study, is the Bayesian one. The basic advantage of this method is that it makes better use of the existing experience of the experimenter and his knowledge of the phenomenon being studied. It aggregates the information prior to the execution of the experiment with the observations after. This different concept uses

Bayes' Theorem, and may result in smaller sample sizes while providing the same sized confidence interval.

In Chapter III, we will describe Bayes' Theorem with the prior, sampling, and posterior distributions, will explain the use of the experimenter's prior bounds on the proportion and the choice of the Uniform distribution as prior, and will give the derivation of the posterior distribution and its properties. Then, in Chapter IV, we will calculate the sample size needed to estimate a proportion and we will compare the results with the classical method. We will explain the computer program used for the Bayesian results and will provide tables and examples to assist the reader. The final chapter will summarize our work, and suggest additional applications of Bayes' Theorem to reduce the cost of weapon system testing.

## II. SAMPLE SIZE TO ESTIMATE A PROPORTION USING THE CLASSICAL METHOD

In this chapter, we will explain the classical method to find the sample size to estimate a proportion. First we will find a *point estimate* of our proportion or probability which is an estimate given by a single number. Then we will find an *interval estimate*, given by two numbers between which our proportion must be considered to lie. Interval estimates provide an indication of the precision or accuracy of an estimate and are therefore preferable to point estimates. Finally, we will use this confidence interval to determine the number of samples needed to achieve a particular interval size.

### A. THE POINT ESTIMATE FOR A PROPORTION

Generally, an estimation problem consists of the manipulations we might make of the observed values in a sample to get a good guess, or estimate of the value of an unknown parameter or parameters.

In our case, we have a sample of n items. The probability of occurrence of an event (detect a defective item), called its success, is p while the probability of non-occurrence of the event is 1 - p. We inspect all the n items and count the number of successes as a sequence

of independent Bernoulli trials. Let $x_i$ be the outcome of each trial, where

$$x_i = 1 \quad \text{if we have a success, and}$$
$$x_i = 0 \quad \text{if otherwise,}$$

and let x be the total number of successes. Then the point estimate for our proportion will be the sample proportion

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x}{n} \qquad (2.1)$$

where x follows a binomial distribution. The distribution of sample proportions has mean $\mu_p$ and standard deviation $\sigma_p$ given by

$$\mu_p = p \quad \text{and} \quad \sigma_p = \sqrt{\frac{p(1-p)}{n}} \qquad (2.2)$$

[Ref. 3:p. 142].

For large values of n, the distribution of sample proportions is approximately normally distributed. In particular;

> The normal curve gives an excellent approximation to the binomial distribution when p is close to 0.5. In fact, for p = 0.5, the approximation is good for n as small as 10. As p deviates from 0.5, the approximation gets worse and worse. On the other hand, for values of p significantly different from 0.5, the approximation of the normal distribution to the binomial distribution gets better, the larger the value of n. Even if p is as low as 0.10 or as high as 0.90, if n runs above 50, the normal approximation does not give bad results. Below 0.10 or above 0.90, the Poisson distribution is commonly used to approximate the binomial distribution, although the normal distribution still does fairly well so long as pn ≥ 5 [Ref. 4:p. 100].

## B. THE CONFIDENCE INTERVAL FOR A PROPORTION

Let $\sigma_s$ be the standard deviation of the sampling distribution of a statistic S. If the sampling distribution is approximately normal, we can expect to find, or we can be ***confident*** of finding, an actual sample statistic S lying in the interval $E[S] - 3\sigma_s$ to $E[S] + 3\sigma_s$ about 99.73% of the times. Because of this we call this interval the 99.73% ***confidence interval*** for estimating E[S]. The end values ($S \pm 3\sigma_s$) are the *confidence limits*. Similarly, $S \pm 1.96\sigma_s$ and $S \pm 2.58\,\sigma_s$ are 95% and 99% confidence limits for S. The percentage confidence is called ***confidence level*** and the numbers 1.96, 2.58, etc..., in the confidence limits are called *confidence coefficients* and are denoted by $z_c$. For this study, we will work with the 95% confidence level, the normal approximation to the Binomial, and the corresponding 1.96 confidence coefficient.

If the statistic $S = \hat{p}$ is the proportion of successes from a sample size n drawn from a binomial population in which the proportion or probability of success is p, the confidence limits for p are [Ref. 4:p. 572]

$$\hat{p} \pm z_c \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad . \qquad (2.3)$$

We can compute the confidence limits of Equation 2.3, using the point estimate for our proportion from Equation 2.1 and so the actual probability will lie in the interval

$$\hat{p} - 1.96 \quad \frac{\hat{p}(1 - \hat{p})}{n} \leq p \leq \hat{p} + 1.96 \quad \frac{\hat{p}(1 - \hat{p})}{n} \qquad (2.4)$$

with a 95% confidence level. For example, if from a population we inspect 30 items and 6 are found defective, we can say that we are 95% sure that the true value of the defective proportion p will lie in the interval calculated from the above Equation, where $\hat{p} = 6/30 = 0.2$,

$$0.2 - 0.14 \leq p \leq 0.2 + 0.14,$$

$$\text{or} \quad 0.06 \leq p \leq 0.34.$$

The interval size is 0.34 - 0.06 = 0.28, and this becomes smaller when the sample size increases.

## C. DETERMINING THE SAMPLE SIZE FROM CONFIDENCE INTERVAL

Let's state our problem again, as it was discussed in Chapter I. How many items must we test so that with a certain level of confidence, we can report the reliability of this type of item. The certain level of confidence will be 95% for this study.

One measure of the effectiveness of a sampling effort is the accuracy of the resulting estimates. In our case of estimating a proportion, accuracy is reflected by the size of the resulting 95% confidence interval, or

$$2(1.96)\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \quad .$$

If the experimenter is willing to specify the size of the confidence interval on p that results from his testing,

then his requirement may serve as a basis for specifying sample size.

Let 2A be the desired 95% confidence interval size. Then our proportion will lie between

$$\hat{p} - A \le p \le \hat{p} + A$$

and the interval size is $\hat{p} - A$ to $\hat{p} + A$ or

$$\hat{p} \pm A.$$

From Equation 2.4 we have

$$A = 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \quad , \tag{2.5}$$

and the sample size n can be determined by solving Equation 2.5 [Ref. 7:p. 247]

$$n = \left(\frac{1.96}{A}\right)^2 \hat{p}(1 - \hat{p}) \quad . \tag{2.6}$$

The sample size increases with the accuracy that we want for our estimate. Better accuracy means smaller interval size, that is, smaller A and thus from Equation 2.6 bigger sample size. Also, the sample size is proportional to the square of the confidence coefficient, which reflects the desired confidence level. Finally, the sample size depends on our guess for the proportion p, before we actually sample from the population. We find the first derivative

of Equation 2.6 to be

$$dn/d\hat{p} = (1.96/A)^2 (1-2\hat{p})$$

and the second

$$d^2 n/d\hat{p}^2 = -2(1.96/A)^2 \hat{p}$$

This is negative; so the value of $\hat{p}$ that makes the first derivative zero maximizes n. This happens for $\hat{p} = 0.5$. Thus, our worst case where we need the maximum sample size, is found to be when we guess that half of the population is defective or that we have 50% chance to detect a defective item. In this case we need to sample $n = (1.96/A)^2 (0.5) (0.5)$ or

$$n = 0.9604/A^2 \tag{2.7}$$

items. The sample size decreases when the probability of success increases from 0.5 to bigger values. Finally, this gives an interpretation of the requirement: the value of 2A is the largest confidence interval that the experimenter is willing to have result from his sampling. The value of n given by Equation 2.7 will guarantee that his requirement is met.

Table 1 shows the required number of samples to obtain different 95% confidence interval sizes for various reliabilities.

TABLE 1. NUMBER OF SAMPLES TO OBTAIN 95% CONFIDENCE INTERVAL SIZE
(Results rounded up)

| Interval Size = 2A | Probability Of Success = p 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 0.975 |
|---|---|---|---|---|---|---|
| 0.05 | 1,537 | 1,476 | 1,291 | 984 | 554 | 150 |
| 0.10 | 385 | 367 | 323 | 246 | 139 | 38 |
| 0.15 | 171 | 164 | 144 | 110 | 62 | 17 |
| 0.20 | 97 | 93 | 81 | 62 | 35 | 10 |
| 0.25 | 62 | 60 | 52 | 40 | 23 | 6 |
| 0.30 | 43 | 41 | 36 | 28 | 16 | 5 |

From the above table, we see that if we think that our probability of success will be p = 0.8 and we want to be ±0.10 accurate with 95% confidence level, we have to sample 62 items.

The numbers of the above table are used to construct the graph in Figure 1. Here, we visualize better the previous discussion about the changes of the sample size because of interval size and probability of success.

In Chapter III, we will solve our problem with the use of Bayesian methods, which give better results, that is, smaller numbers of samples than those in Table 1.

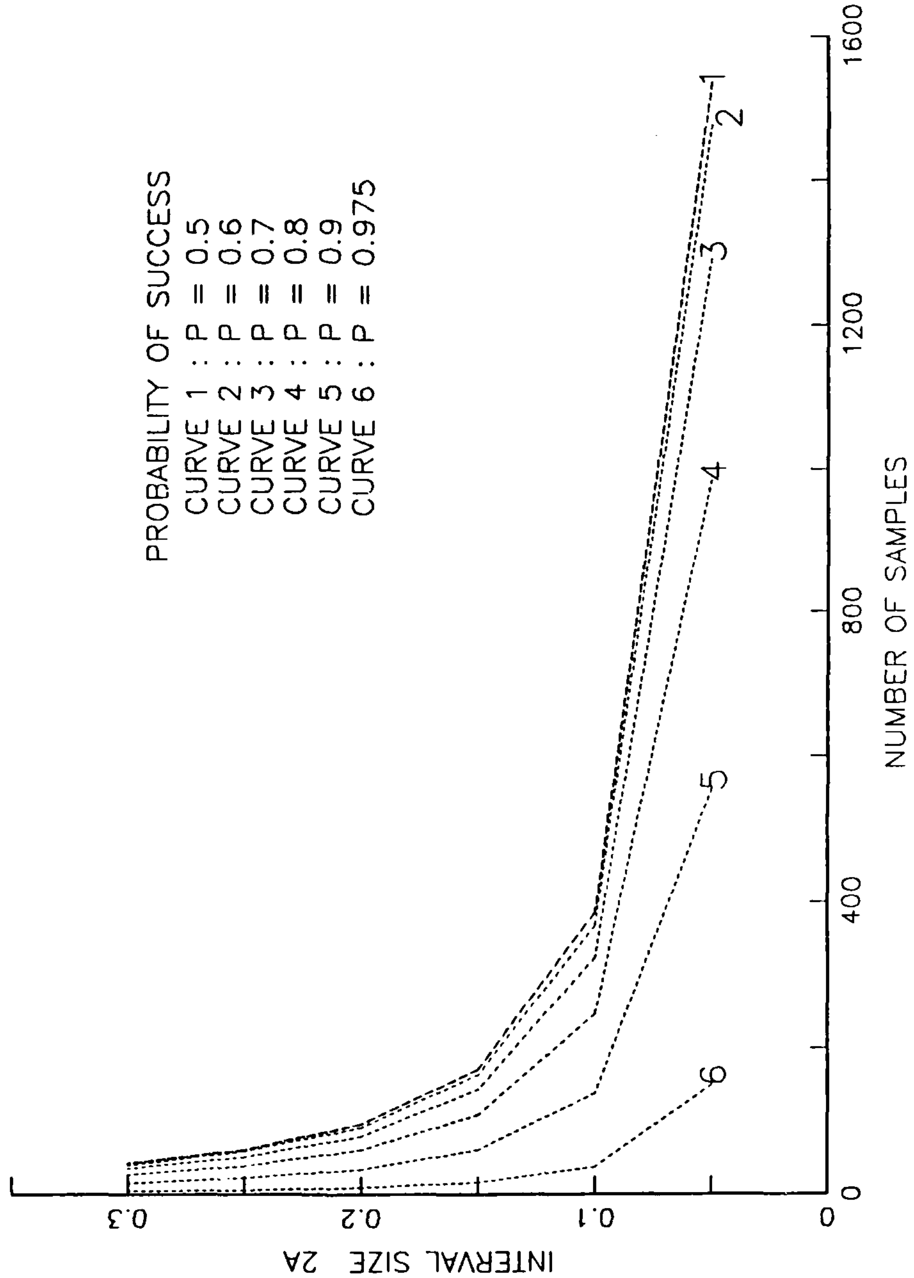


**Figure 1. Number of Samples vs Interval Size For Various Probabilities of Success, With 95% Confidence Interval**

# III. THE BAYESIAN METHOD AND ESTIMATORS

In this chapter, we will explain a Bayesian method to find the sample size to estimate a proportion. To do this, we will first recall Bayes' Theorem and we will use it in our problem. Then we will explain analytically the three parts of the Bayesian result we found: the prior distribution, the sampling, and the posterior one. We will state our reasons for the selection of the Uniform in the interval [a,b] as prior and Binomial as sampling distribution. After that, we will derive the posterior distribution and its first two moments and we will find the Bayes estimators. Finally, we will explain the assumption we made in order to use the posterior distribution to calculate the 95% confidence interval of our proportion.

Inferentially, the Bayesian method permits the use of the knowledge and past experience of the experimenter, before observations are taken. Those, in combination with the sampling results, may give a smaller number of samples to estimate a proportion than that given by the classical method.

## A. BAYES' THEOREM

One different method to estimate a proportion is to use Bayes' Theorem. Let us explain the procedure stating Bayes' Theorem first.

Suppose that $A_1$, $A_2, \ldots,$ $A_n$ are mutually exclusive events whose union is the sample space S, i.e., they form a *partition* of the S and one of them must occur. Then if A is any event of S, we have the following ***Bayes' Theorem:***

$$P(A_k/A) = \frac{P(A_k)\ P(A/A_k)}{\sum_{k=1}^{n} P(A_k)\ P(A/A_k)} \quad . \tag{3.1}$$

Consider now our problem. If a lot has a defective proportion p, then the probability that a sample of size n will contain exactly X defective items is, for relatively large lots, approximately [Ref. 4:p. 558],

$$P(X/p) = \binom{n}{X} p^x \ (1-p)^{n-x} \quad . \tag{3.2}$$

Suppose now that p is itself a continuous random variable with density function f(p), where

$$\int_{-\infty}^{\infty} f(p)dp = 1 \quad .$$

Then the joint probability that for a given lot, (1) p will fall in the interval p to p + dp and, (2) that a sample size n taken from this lot contains X defective items, is the product

$$P(X,p) = P(X/p)\ f(p) \quad .$$

According to Bayes' Theorem above, for the continuous case, the probability that the p that produced the given X lies in the interval p to p + dp is

$$P(p/X) = \frac{P(X/p)\ f(p)}{\int_0^1 P(X/p)\ f(p)dp} \quad . \tag{3.3}$$

The density function f(p) is called the *prior probability distribution* and the probability P(p/X) is the *posterior probability distribution*. The third part of the above Equation 3.3, the sampling distribution P(X/p), is the probability function from which we will take the X items. Because we count successes in repeated n Bernoulli trials, this is Binomial as in Equation 3.2.

## B. SELECTION OF THE PRIOR DISTRIBUTION

The prior distribution of a parameter p is a probability function or probability expressing our degree of belief about the value of p, prior to observing a sample of a random variable X whose distribution depends on p [Ref. 2:p. 553]. In other words, we can assign a prior distribution to a parameter p when we have enough information about the relative frequencies with which p has taken each of its possible values in the past. For example, suppose that the proportion p of defective items in a certain lot is unknown. Suppose also that this lot is made from a manufacturer who has produced many such lots in the past and that detailed records have been kept about the defective fractions in these lots. The relative frequencies for these past lots can be used to estimate a prior distribution for p, which can be used in our certain lot.

Different distribution functions can be characterized as "priors". As examples, for a bounded variable p, we

mention the Uniform distribution on the interval [0,1], a triangular shaped distribution, and the Beta distribution with various parameter values. The Beta distribution for $0 \le p \le 1$ was used as the prior in Ref. 6, where the sample size problem for a Bayesian confidence interval was also addressed.

On the other hand, we must note that the prior distribution "is a subjective probability distribution in the sense that it represents an individual experimenter's information and subjective beliefs about where the true value of p is likely to lie". [Ref. 5:p. 314] Often the best prior information about the parameter p may simply be bounds on p, wherein the experimenter can only say that p will not exceed some value b, and will not be less than some value a. The density function that is reasonable to combine with experience expressed as bounds on the unknown parameter seems to be the ***Uniform distribution on the interval [a,b]*** since it "distributes our ignorance equally" in the prior known interval [Ref. 4:p. 560].

The Uniform density function and prior distribution for this study is the Uniform [a,b]

$$f_1(p) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le p \le b \\ 0 & \text{otherwise,} \end{cases} \tag{3.4}$$

where

$$0 \le a \le p \le b \le 1 \quad .$$

Note that the Uniform [0,1] distribution belongs to the class of Beta (r,s) distributions when both Beta parameters are 0.

It is valuable to remember here that the Beta density function (in the form that we use extensively later) is

$$f(x) = \begin{cases} \dfrac{(r+s+2)}{(r+1)\quad(s+1)}\, x^r\ (1-x)^s & \text{for } 0 \le x \le 1 \\ \\ 0 & \text{otherwise,} \end{cases}$$

where

$$r > -1 \text{ and } s > -1 \quad .$$

C. DERIVATION OF THE POSTERIOR

The posterior density as it is expressed in Equation 3.3, Bayes' Theorem, is simply the conditional density of p, given the sample values. It "expresses our degree of belief of the location of p, given the results of the sample". [Ref. 2 p. 556]

To derive our posterior distribution $f_2(p/x)$, we substitute into Equation 3.3 the prior as the Uniform density from Equation 3.4, and the sampling distribution as the Binomial from Equation 3.2:

$$f_2(p/x) = \frac{\binom{n}{x}\ p^x(1-p)^{n-x}\ \dfrac{1}{b-a}}{\int_a^b \binom{n}{x}\ p^x(1-p)^{n-x}\ \dfrac{1}{b-a}\,dp} \quad ,$$

where

$$0 \leq a \leq p \leq b \leq 1 \quad .$$

If we cancel out terms, we have

$$f_2(p/x) = \frac{p^x (1 - p)^{n-x}}{\int_a^b p^x (1 - p)^{n-x} \, dp} \quad .$$

If we multiply numerator and denominator by the same number, we have

$$f_2(p/x) = \frac{\dfrac{\Gamma(n+2)}{\Gamma(x+1)\ \Gamma(n-x+1)} \ p^x (1-p)^{n-x}}{\int_a^b \dfrac{\Gamma(n+2)}{\Gamma(x+1)\ \Gamma(n-x+1)} \ p^x (1-p)^{n-x} dp} \quad ,$$

and we notice that the denominator is the area under the curve of a Beta distributed random variable with parameters

$$r = x + 1$$

and $\quad s = n - x + 1 \quad ,$

over the interval a to b. Thus, we have

$$f_2(p/x) = \frac{\dfrac{\Gamma(n+2)}{\Gamma(x+1)\ \Gamma(n-x+1)} \ p^x (1-p)^{n-x}}{F_3(b) - F_3(a)} \quad ,$$

where $F_3$ is the CDF of a Beta (r,s) distribution.

We also note that the numerator has the form of a Beta density function with the same parameters (r,s) but the argument p is defined to be $a \leq p \leq b$ .

So finally, our posterior distribution becomes

$$f_2(p/x) = \frac{f_3(p)}{F_3(b) - F_3(a)} \tag{3.5}$$

where $0 \le a \le p \le b \le 1$ ,

$F_3$ : CDF of Beta $(r,s)$

and $f_3(p)$ has the form of Beta $(r,s)$ ,

where $r = x + 1$

and $s = n - x + 1$ .

If we let $c = 1/[F_3(b) - F_3(a)]$, this posterior distribution has the functional form of a Beta density function with parameters $(r,s)$, multiplied by a positive constant $c \ge 1.0$, for a random variable p bounded by the bounds $(a,b)$ of the Uniform prior distribution.

Let us illustrate with an example using the above conclusions. Suppose that an experimenter uses past data and puts bounds $a = 0.2$ and $b = 0.8$ on the probability that a defective item is located in a sample of $n = 10$ items. His prior distribution is Uniform for a random variable p bounded between 0.2 and 0.8. After the inspection of all the items, he counts $x = 5$ defective. Then from Equation 3.5, we have $r = 5 + 1 = 6$ and $s = 10 - 5 + 1 = 6$. We also have from Beta CDF that $F_3(0.8) = 0.98834$, $F_3(0.2) = 0.01165$ and $c = 1/[F_3(0.8) - F_3(0.2)] = 1/0.97669 = 1.02386$. Then, his posterior distribution, is a form of

Beta (6,6) multiplied by 1.02386, for a random variable p bounded again between 0.2 and 0.8.

## D. THE POSTERIOR DISTRIBUTION AND BAYES' ESTIMATORS

We have shown above an example of the density function of the posterior distribution, for specified Uniform prior and Binomial sampling distributions. We will now calculate the mean and the variance of the posterior distribution in the general case of Equation 3.5. Let c be the constant factor $1/[F_3(b)-F_3(a)]$ in the Equation 3.5. Then

$$E[p|x] = \int_a^b pc \frac{\Gamma(n+2)}{\Gamma(x+1)\,\Gamma(n-x+1)} p^x (1-p)^{n-x}\, dp \quad .$$

If we combine terms and multiply numerator and denominator by n + 2 and x + 1 respectively, we have

$$E[p|x] = c\,\frac{x+1}{n+2} \int_a^b \frac{\Gamma(n+3)}{\Gamma(x+2)\,\Gamma(n-x+1)} p^{x+1} (1-p)^{n-x}\, dp$$

$$= c\,\frac{x+1}{n+2} \int_a^b f_4\,(p)\, dp \quad ,$$

where $f_4$ a form of Beta ($r = x + 2$, $s = n - x + 1$) and $a \le p \le b$.

Substituting for c, we have

$$E[p|x] = \frac{x+1}{n+2}\;\frac{F_4(b) - F_4(a)}{F_3(b) - F_3(a)} \quad , \tag{3.6}$$

where $F_4$ : CDF of Beta ($r = x + 2$, $s = n - x + 1$)

and $F_3$ : CDF of Beta ($r = x + 1$, $s = n - x + 1$).

It is worthwhile to give a different form of the mean. From Equation 3.6, we have

$$E[p|x] = \frac{x+1}{n+2} \frac{\int_a^b \frac{\Gamma(n+3)}{\Gamma(x+2)\ \Gamma(n-x+1)}\ p^{x+1}\ (1-p)^{n-x}\ dp}{\int_a^b \frac{\Gamma(n+2)}{\Gamma(x+1)\ \Gamma(n-x+1)}\ p^{x}\ (1-p)^{n-x}\ dp}$$

Since the arguments of the Gamma functions are integers, we can substitute with factorials and pull them out of the integrals, giving

$$E[p|x] = \frac{x+1}{n+2}\ \frac{\frac{(n+2)!}{(x+1)!\ (n-x)!}\ \int_a^b p^{x+1}(1-p)^{n-x}dp}{\frac{(n+1)!}{x!\ (n-x)!}\ \int_a^b p^{x}(1-p)^{n-x}dp}\ .$$

Simplifing the constant terms, we have

$$E[p|x] = \frac{(n+2)!\ x!\ (x+1)\ (n-x)!\ \int_a^b p^{x+1}(1-p)^{n-x}dp}{(n+1)!(n+2)(x+1)!(n-x)!\ \int_a^b p^{x}(1-p)^{n-x}dp}$$

or

$$E[p|x] = \frac{(n+2)!\ (x+1)!\ (n-x)!}{(n+2)!\ (x+1)!\ (n-x)!}\ \frac{\int_a^b p^{x+1}(1-p)^{n-x}dp}{\int_a^b p^{x}(1-p)^{n-x}dp}$$

which obviously gives

$$E[p|x] = \frac{\int_a^b p^{x+1}(1-p)^{n-x}dp}{\int_a^b p^{x}(1-p)^{n-x}dp}\ .$$

To calculate the Variance, we use the well known result $Var[p] = E[p^2] - E[p]^2$. Working as we did for the mean, we have that

$$E[p^2|x] = c\,\frac{x+1}{n+2}\,\frac{x+2}{n+3}\,(F_5(b) - F_5(a)) \quad ,$$

where $F_5$ is a Beta CDF with parameters $r = x + 3$ and $s = n - x + 1$. Then the variance of the posterior distribution is

$$Var(p|x) = c\,\frac{x+1}{n+2}\,\frac{x+2}{n+3}\,(F_5(b) - F_5(a))$$

$$- \; c\,\frac{x+1}{n+2}\,(F_4(b) - F_4(a))^2 \quad ,$$

or

$$Var(p|x) = \frac{x+1}{n+2}\,\frac{x+2}{n+3}\,\frac{F_5(b) - F_5(a)}{F_3(b) - F_3(a)}$$

$$-\left(\frac{x+1}{n+2}\,\frac{F_4(b) - F_4(a)}{F_3(b) - F_2(a)}\right)^2 \quad . \qquad (3.7)$$

The commonly used point estimate $\hat{p}$ for a proportion in the Bayesian method is the mean of the posterior distribution. However, if the posterior density for p is not symmetric, other measures of the middle of the posterior might also be used as the pcint estimate. Two such measures are the mode of the posterior (which maximizes the posterior density) or the median (which is

the value that splits equally the area under the density curve).

As in the classical method, interval estimates are preferable to point estimates. In this Bayesian method, it is easy to construct interval estimates of the proportion p. A 95% confidence interval is provided by the 2.5 and 97.5 quantiles of our posterior distribution of Equation 3.5. Thus, the interval estimates depend on the subjective bounds of the prior Uniform distribution (a and b), the sample size n, and the number x of successes from the sampling (Binomial distribution for sequential Bernoulli trials). Before sampling, we know a and b, but we do not know n and x. Note that our problem is how big should be the sample size n.

This results in a need to guess the number x of successes before we actually sample. We recall from the definition of the mean of a random variable, that this number would locate the center of gravity of the distribution of the random variable and thus, is a likely candidate if we have to give a single number as our guess of the value of the random variable. So, prior to sampling, a "good" guess for x, the number of successes, is the mean of the prior Uniform distribution multiplied by the sample size n, or

$$x = \left(\frac{a + b}{2}\right) n \quad .$$

Using this, our posterior distribution in Equation 3.5 becomes, for purposes of determining sample size n,

$$f_2(p/x) = \frac{f_3(p)}{F_3(b) - F_3(a)} \quad , \qquad (3.8)$$

where $0 \le a \le p \le b \le 1$ ,

$F_3$ : CDF of Beta $(r^*, s^*)$

and $f_3(p)$ has the form of Beta $(r^*, s^*)$ ,

where $r^* = \left(\frac{a+b}{2}\right) n + 1$

and $s^* = n - \left(\frac{a+b}{2}\right) n + 1$ .

In Equation 3.8, the posterior distribution has the functional form of a Beta density function with parameters $(r^*, s^*)$, multiplied by a positive constant $c \ge 1.0$, for a random variable p bounded by the bounds (a,b) of the prior Uniform distribution.

From now on Equation 3.8 will be our posterior distribution which we use in the next chapter as the base to develop a procedure to calculate the interval estimates and sample size. In the next chapter also, we will discuss the computer programs that we used to find the sample sizes to estimate the proportion p.

# IV. SAMPLE SIZE TO ESTIMATE A PROPORTION USING THE BAYESIAN METHOD

In this chapter, we will explain how to use the Bayesian method with a Uniform prior in order to find the sample size to estimate a proportion. The experimenter has some prior information about the unknown proportion in the form of upper and lower bounds on the unknown proportion. We use them as bounds for a prior Uniform distribution and we wish to determine the sample size he needs, based on the accuracy he likes.

First, we will derive the bounds of the Bayesian confidence interval, where the proportion to be estimated should lie with 95% confidence level. This interval leads to the necessary sample size. Then, we will discuss the computer programs we have used in this procedure. Finally, we will provide tables and examples to assist the user to find the sample size that meets his goals, and to visualize the advantage of the Bayesian method in giving smaller samples than the classical one.

## A. THE BAYESIAN CONFIDENCE INTERVAL

Once we have obtained our posterior distribution, we can construct an interval which contains $100(1-\alpha)\%$ of the posterior probability. Our posterior distribution, given by Equation 3.8, has the Beta $(r^*, s^*)$ form, multiplied by a

constant for a random variable p bounded between a and b. We let $\alpha$ be 0.05 for this study, and thus want a 95% confidence interval size.

Let p.lo and p.up be the lower and upper bounds of the confidence interval. Letting the area between a and p.lo, the lower bounds, be $\alpha/2 = 0.025$ of the whole area under our posterior density function, $f_2(p/x)$, we have

$$F_2(p.lo) = a/2 = 0.025 \quad .$$

From Equation 3.8, we have also that

$$F_2(p.lo) = \int_a^{p.lo} f_2(p/x)\,dp,$$

or

$$F_2(p.lo) = \frac{1}{F_3(b) - F_3(a)} \int_a^{p.lo} f_3(p)\,dp,$$

where $F_3(p)$ is Beta $(r^*, s^*)$.

Thus, we have the equation

$$\frac{1}{F_3(b) - F_3(a)} \int_a^{p.lo} f_3(p)dp = 0.025,$$

or

$$F_3(p.lo) - F_3(a) = 0.025\,[F_3(b) - F_3(a)],$$

which finally gives

$$F_3(p.lo) = 0.025\,F_3(b) + 0.975\,F_3(a) \quad . \qquad (4.1)$$

Similarly, letting $\alpha/2 = 0.025$ be the area under the posterior density function $f_2(p/x)$ between p.up and b, we have

$$F_3(p.up) = 0.0975\ F_3(b) + 0.025\ F_3(a) \quad . \qquad (4.2)$$

From Equations 4.1 and 4.2, we calculate p.lo and p.up, and then by subtracting p.lo from p.up, we get the 95% confidence interval size. This is the ***Bayesian interval*** where our proportion to be estimated should lie 95% of the time. As we did with classical method, let us call the size of this interval 2A; it is a measure of estimation accuracy. We will see that the sample size depends upon the prior bounds (a,b) and upon the interval size 2A. The decision maker uses his past information to state the bounds and his preference in accuracy to state the interval size. The Bayesian method of this study gives the interval p.lo to p.up, where the proportion p lies with 95% confidence level.

In the next section, we will explain how these values may be used to find the required sample size.

## B. DETERMINING THE SAMPLE SIZE FROM THE BAYESIAN INTERVAL

In Chapter III, we derived our posterior density function which has the form of a Beta distribution with parameters $(r^*, s^*)$, multiplied by a constant, i.e.,

$$f_2(p/x) = \frac{f_3(p)}{F_3(b) - F_3(a)}$$

where $0 \le a \le p \le b \le 1$,

$F_3$: CDF of Beta $(r^*, s^*)$

and $f_3(p)$ has the form of Beta $(r^*, s^*)$.

In Chapter III, we also explained why, for purposes of determining the sample size n, the parameters $r^*$ and $s^*$ take the values

$$r^* = \left(\frac{a+b}{2}\right)n + 1 \quad ,$$

and

$$s^* = n - \left(\frac{a+b}{2}\right)n + 1 \quad ,$$

where a and b are the bounds of the prior Uniform distribution.

Once we have obtained the parameters of our posterior distribution, using Equations 4.1 and 4.2 we need to compute the inverse cumulative distribution function at 0.025 and 0.975 for a Beta with parameters $r^*$ and $s^*$. This will result in the lower and upper bounds of the 95% confidence interval. Then, if we subtract the lower from the upper bound, we can determine the size of the desired confidence interval.

The above procedure is used in the APL program SAMPLE located in Appendix A. This program computes the sample

size needed to obtain a 95% confidence interval where the probability or proportion should lie. The program is interactive and requires the user to input the bounds of the prior Uniform distribution and the desired confidence interval size. Then it calculates the parameters of the posterior distribution, and computes the confidence interval that is provided when we sample 10 items. Consequently, using a loop, it increases the sample size until the desired confidence interval size is reached. Finally, it prints the 95% confidence interval bounds and the sample size needed to obtain the required confidence interval.

The program SAMPLE uses the subroutines BQUAN, NQUAN, and BETA located in Appendix B. These are APL programs designed at Naval Postgraduate School to compute the inverse cumulative distribution function of Beta distribution. It must be noted that BQUAN often cannot compute the inverse cumulative distribution function for large Beta parameters. In our case, large parameter values mean large sample size. Thus SAMPLE was written to terminate its calculations at sample size 150: results can always be obtained for sample sizes at or below this value. This number can be increased, but in general SAMPLE cannot evaluate sample sizes greater than say 200.

## C. TABLES FOR FINDING THE SAMPLE SIZE

In this section, we provide a table to assist the user in finding the sample size that reflects his past knowledge with the prior bounds, and his preference for accuracy with the 95% confidence interval size.

Table 2 was constructed by executing the APL program SAMPLE repeatedly for selected values of a and b, covering the whole range from 0 to 1.0. Also, convenient interval sizes were used.

The use of Table 2 is simple. For example, suppose the user puts prior bounds 0.5 to 0.8 and wants the interval size 2A to be 0.25. He looks at the part of the table for b = 0.8 and he finds the entry in row a = 0.5 and column CI size 2A = 0.25. He has to sample 40 items to be 95% confident that the proportion p will be in a confidence interval of size 0.25.

Before giving more examples of the use of this table, it is well to note that some entries in Table 2 are blank. One reason for missing entries is, as mentioned in the previous section, that SAMPLE generally can not evaluate sample sizes greater than 200. Another problem in constructing Table 2 occurred with the APL program NQUAN. Its execution stops when the sample size is big (greater than 100) and the sum of the prior bounds (a + b) is between 0.7 and 1.3. Finally, another reason for blank entries is that 2A must be less than b - a.

TABLE 2. NUMBER OF SAMPLES TO OBTAIN 95% CONFIDENCE BAYESIAN INTERVAL

| | b = .05 | b = .1 | | b = .2 | | |
|---|---|---|---|---|---|---|
| a | CI size 2A 0.025 | CI size 2A 0.05 | 0.075 | CI size 2A 0.075 | 0.10 | 0.15 |
| 0 | 617 | 299 | 82 | 249 | 140 | 40 |
| 0.025 | | 336 | | 275 | 151 | 32 |
| 0.05 | | | | 299 | 156 | |
| 0.075 | | | | 318 | 139 | |

| | b = 0.3 | | | | | b = 0.4 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | CI size 2A | | | | | CI size 2A | | | | |
| a | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 |
| 0 | 196 | 87 | 42 | 12 | | 245 | 108 | 60 | 36 | 18 |
| 0.05 | 221 | 94 | 37 | | | | 118 | 64 | 35 | 15 |
| 0.1 | 244 | 93 | | | | | 125 | 66 | 30 | |
| 0.2 | | | | | | | 126 | | | |

| a | b = 0.5 | | | | b = 0.6 | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 126 | 70 | 44 | 29 | 141 | 78 | 49 | 29 |
| 0.05 | 134 | 74 | 46 | 30 | 147 | 82 | 51 | 35 |
| 0.1 | 141 | 78 | 48 | 29 | 153 | 85 | 53 | 36 |
| 0.2 | 153 | 82 | 40 | | | 89 | 55 | 35 |
| 0.3 | | | | | | 90 | 44 | |

(TABLE 2 CONTINUED)

| a | b = 0.7 | | | | | b = 0.8 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | CI size 2A | | | | | CI size 2A | | | | |
| | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 |
| 0 | | 153 | 85 | 53 | 36 | | | 89 | 56 | 38 |
| 0.05 | | | 87 | 55 | 37 | | | 91 | 57 | 39 |
| 0.1 | | | 89 | 56 | 39 | | | 92 | 58 | 39 |
| 0.2 | | | 92 | 58 | 39 | | 168 | 93 | 59 | 40 |
| 0.3 | | 168 | 93 | 58 | 36 | | | 92 | 58 | 39 |
| 0.4 | | | 90 | 44 | | | | 89 | 55 | 35 |
| 0.5 | | | | | | | 153 | 82 | 40 | |
| 0.6 | | | | | | | 126 | | | |

| a | b = 0.9 | | | | | b = 0.95 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | | 92 | 58 | 39 | | | 93 | 58 | 40 |
| 0.05 | | | 93 | 58 | 40 | | | 93 | 59 | 40 |
| 0.1 | | 168 | 93 | 59 | 40 | | | 93 | 58 | 40 |
| 0.2 | | | 92 | 58 | 39 | | | 91 | 57 | 39 |
| 0.3 | | | 89 | 56 | 38 | | | 87 | 55 | 37 |
| 0.4 | | 153 | 85 | 53 | 36 | | 147 | 82 | 51 | 35 |
| 0.5 | | 141 | 78 | 48 | 29 | | 134 | 74 | 46 | 30 |
| 0.6 | | 126 | 66 | 30 | | | 118 | 64 | 35 | 15 |
| 0.7 | 244 | 93 | | | | 221 | 95 | 37 | | |
| 0.8 | | | | | | 156 | | | | |

(TABLE 2 CONTINUED)

| a | b = 0.975<br>CI size 2A<br>0.05 | b = 1.0<br>CI size 2A<br>0.05 |
|---|---|---|
| 0.85 | 495 | 432 |
| 0.9 | 336 | 299 |
| 0.925 | | 192 |

| | b = 0.975<br>CI size 2A | | | | | b = 1.0<br>CI size 2A | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| a | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 |
| 0 | | 168 | 93 | 58 | 40 | | 168 | 93 | 59 | 40 |
| 0.05 | | | 93 | 58 | 40 | | | 93 | 58 | 40 |
| 0.1 | | | 93 | 58 | 40 | | | 92 | 58 | 39 |
| 0.2 | | | 90 | 57 | 39 | | | 89 | 56 | 38 |
| 0.3 | | 152 | 86 | 54 | 37 | | 153 | 85 | 53 | 36 |
| 0.4 | | 144 | 80 | 50 | 34 | | 141 | 78 | 49 | 29 |
| 0.5 | | 130 | 72 | 45 | 29 | 288 | 125 | 70 | 43 | 29 |
| 0.6 | 256 | 113 | 62 | 36 | 17 | 245 | 108 | 60 | 36 | 18 |
| 0.7 | 209 | 91 | 41 | 8 | | 196 | 87 | 42 | 12 | |
| 0.75 | 183 | 72 | 15 | | | 169 | 71 | 21 | | |
| 0.8 | 151 | 32 | | | | 140 | 40 | | | |
| 0.85 | 81 | | | | | 92 | | | | |

We should mention also that all the programs used in this study are written in APL and can be run on any computer with APL capabilities. However, because of the extensive loops they have, they may require a significant amount of time.

Let us illustrate with some additional examples the way that Table 2 can be used, and at the same time, let us compare the results it gives with those given in Table 1 from classical statistics.

The classical procedure requires the experimenter to state values for confidence interval size 2A and estimated probability of success p. The Bayesian procedure requires 2A and prior bounds a, b. In order to be able to compare the results obtained from these two methods, we recall an argument in which our Bayesian procedure was based. The experimenter, using his past knowledge, states the bounds of the prior Uniform distribution.

> For any finite sample size, the Bayesian estimate is "shaded" toward the prior mean, the best guess for $\theta$ before any sample values were taken [Ref. 2:p. 566].

In our study, we use p in place of $\theta$. The mean of the prior is $(a + b)/2$, the sum of the bounds divided by two. To compare the two methods, this number, from the Bayesian method, is used as the probability of success to enter Table 1 and find the suggested sample size from the classical method. Thus, if the bounds of the prior Uniform are 0.5 to 0.9, we use the number $(0.5 + 0.9)/2 = 0.7$ as

the probability of success to enter in Table 1, and then compare the results.

We proceed now with some examples to explain the use of the Bayesian Tables and how to find the sample size when we have the bounds of the prior Uniform distribution and the desired 95% confidence interval size.

1. Example 1: Fraction Defective

   Suppose a lot of 10,000 items is received from a supplier; the lot contains p (unknown) defective items. Also, suppose that we have kept records for the past lots from this supplier and we decide our subjective bounds to be $a = 0$ and $b = 0.4$. How many items do we have to sample in order to be 95% confident for the fraction defective, with $\pm$ 0.01 estimation accuracy? We look at the part of Table 2 with $b = 0.4$, and we find the sample size of 60 items in the entry for row $a = 0$ and column CI size $2A = 0.2$. If we look at Table 1, for probability of success 0.2 (column 0.8) and interval size 0.2, we find 62 items. The Bayesian approach reduced a sample size by 2/62 or 3%. Note also that had we used the common textbook formula for sample size for a proportion, (Equation 2.7), the result is $n = 97$.

2. Example 2: Hit Probability

   Suppose the size of the load of a recently modified weapon system has to be decided; this system has (unknown) hit probability, p, against one of the targets it is designed for. Suppose also that we have data from the past firings with the old version and we decide our subjective bounds will be $a = 0.7$ and $b = 0.9$. How many items do we have to fire in order to be 95% confident of the hit probability, with $\pm$ 0.05 estimation accuracy? We look at the part of Table 2 for $b = 0.9$ and we find the number of 244 items in the entry for row $a = 0.7$ and column CI size $2A = 0.1$. For the classical result, we look at Table 1, for $\hat{p} = 0.8$ and $2A = 0.1$, we find 246 items. With the Bayesian approach, we need 2 items less, i.e., 0.008% better results. Had we used Equation 2.7, the result would be 385.

3. Example 3: Detection Probability

Suppose the number of acoustic devices has to be defined in order to design a new type of sonobuoy. Each device has p (unknown) probability of detection. Suppose also that experience from similar hydrophones gives a probability between 0.1 and 0.5, and we like ± 0.15 accuracy. How many acoustic devices will we have to use in order to fulfill the requirements 95% of the time? We look at the part of Table 2 for b = 0.5 and we find 29 devices in the entry for row a = 0.1 and CI size 2A = 0.30. If we use Table 1 from classical statistics for probability of success 0.3 (column 0.7), we find 36 devices. The Bayesian approach gives 7 devices less, that is 19.5% better results. Table 1 in column 0.5 (Equation 2.7) gives the number 43, almost twice as big as the Bayesian one.

We have demonstrated above the use of the Bayesian tables in finding sample sizes. The tables provide the most common bounds and interval sizes. If the user needs a sample size for bounds and/or interval size that are not included in the tables, again the program SAMPLE can be used. The user interactively inputs his prior bounds and desired 95% confidence interval size and the output is the confidence interval bounds and the number of samples. An APL session, solving a problem in this case, is shown in Figure 2. Prior bounds are 0.40 to 0.75 and interval size is 0.25. The Bayesian method requires 54 items .

In order to compare the answer with that from classical method, we use the Equation 2.6 for $\hat{p}$ = (0.40 + 0.75)/2 = 0.575 and A = 0.25/2 = 0.125 and we find 61 items. In this case, the Bayesian approach gave 7 items less or 7/61 = 11.5% better results.

To assist the comparisons between the two methods, let us present the results of the examples in a table.

```
      SAMPLE
ENTER A, LOWER BOUND OF UNIFORM PRIOR
⎕:
      .4

ENTER B, UPPER BOUND
⎕:
      .75

ENTER 95 PERCENT C.L. INTERVAL SIZE (MUST BE LESS THAN B-A)
⎕:
      .25

95 PERCENT C.L. UPPER BOUND :  0.6959232423
      >>        LOWER       :  0.4467326643

      >>       INTERVAL SIZE:  0.2491905779

REQUIRED SAMPLE SIZE  :  54
```

Figure 2. An APL Session Using The Program SAMPLE

TABLE 3. NUMBER OF SAMPLES TO OBTAIN 95% CONFIDENCE INTERVAL SIZE

| Example | a | b | CI Size | Number of Samples Bayesian | Classical | Improvement % |
|---|---|---|---|---|---|---|
| Example 3 | 0.1 | 0.5 | 0.3 | 29 | 36 | 19.5 |
| APL session | 0.4 | 0.75 | 0.25 | 54 | 61 | 11.5 |
| Example 1 | 0.0 | 0.4 | 0.2 | 60 | 62 | 3.0 |
| Example 2 | 0.7 | 0.9 | 0.1 | 244 | 246 | 0.008 |

We see that as the sample size gets larger, the % improvement of the Bayesian method decreases. "The difference between the Bayesian values, and the classical approach, disappears as n increases" [Ref. 2:p. 573]. This happens because as the sample size gets larger, the posterior distribution becomes less dependent on the assumed prior and more on the sampling one. When values of n are smaller, the Bayesian values may differ considerably from the classical. This situation underlines the importance of the prior distribution in that, for small sample sizes, the prior distribution must be chosen carefully.

To complete the study of the Bayesian method, let us look again closely at Table 2. We see that the sample size n increases when the interval size 2A decreases. We see also, that for the same interval size and holding one bound fixed, the sample size increases as the other bound approaches 0.5. If we use the interpretation which we did before, i.e., to consider that the sum of bounds divided by two in Bayesian method is equivalent with the probability of success in the classical method, we conclude that Bayesian intervals behave exactly in the same way with those in classical statistics. The discussion in Chapter II and its presentation with Figure 4, about the changes and the dependence of sample size in classical method, is also valid for the Bayesian one.

In the next chapter, we summarize our work and propose additional studies for the use of Bayesian methods to reduce the sample size and thus, the cost of weapon testing.

# V. SUMMARY AND SUGGESTIONS FOR FURTHER STUDY

In this chapter, we will summarize the procedure we used, working with the Bayesian method, to obtain sample sizes which are smaller than those given by classical statistics. Finally, we make some recommendations for further research in using Bayesian methods to reduce the sample size needed to estimate a proportion or probability in any test field, and thus to reduce the cost.

## A. SUMMARY

In this paper, we used a Bayesian method to obtain the number of samples needed to estimate a proportion or probability.

First, we described the classical method and explained the point and interval estimates. Using desired confidence interval sizes, we produced a table with sample sizes given from classical statistics for 95% confidence intervals.

Then, we described the Bayes' Theorem with the prior, sampling and posterior distributions. We choose the Uniform [a,b] as prior. This, combined with the sampling Binomial, give a form of Beta distribution as posterior for a random variable bounded by [a,b]. We derived the Bayesian 95% confidence interval and produced a computer program to calculate the sample size. We provided a table and gave some examples to assist the user to determine how

to use these results to obtain smaller sample sizes. Finally, we compared the results given by both methods for the same 95% confidence interval size. For small sample sizes, generally smaller than 100, the Bayesian method with the Uniform [a,b] prior improves the results and thus decreases the cost of tests based on sequential Bernoulli trials.

Thus, when the decision maker has prior knowledge, and he wants to benefit from this, the Bayesian method of this study is recommended in order to reduce the number of items, and consequently, the cost.

In the next section we suggest some additional studies based on Bayes' Theorem and on this paper, for even smaller sample sizes.

## B. SUGGESTIONS FOR FURTHER STUDY

This paper uses the Uniform [a,b] distribution as prior. This prior is easy to use, but it does not always give better answers than other prior distributions. The study by Manion approached the sample size question with a Beta prior distribution for a proportion bounded by 0 and 1.0. For a quick comparison, we use an example.

> As an example, if the decision maker wanted the size of the 95% confidence interval to be 0.20 and his subjective bounds on the proportion were 0.14 to 0.86, the parameters on the Beta prior would be 4,4 and the number of observations needed would be 87 [Ref. 6:p. 42].

Our study with Uniform [0.14,0.86] prior gives 93 items. Possibly, it could be better if a prior that

combines both studies and concepts were to be chosen, i.e., a Beta for a prior bounded random variable. Another prior density function could be the triangular one, again for bounded random variable. The sensitivity of resulting sample size to the choice of bounds a and b could also be explored.

An additional research task could be an effort to fill the blanks in the table of the Bayesian interval sample sizes of this study. This presupposes the development of a computer program that can compute the inverse cumulative density function of the Beta distribution for large parameters.

Finally, an addition to this paper could be the development of tables for confidence intervals other than 95%, such as 90%, 97.5%, and 99%.

We hope that the chance to reduce the cost of sampling with smaller sample sizes to estimate a proportion, as given from this paper, will be beneficial to any authority dealing with tests of acceptance, reliability, etc.

## APPENDIX A. THE APL PROGRAM "SAMPLE" USED TO COMPUTE SAMPLE SIZES FOR BAYESIAN INTERVALS WITH 95% CONFIDENCE LEVEL.

```
      ∇ SAMPLE [⎕]
      ∇ SAMPLE;LF;RT;N;C;D;LO;UP;E
[1]   ⍝ THIS PROGRAM COMPUTES THE SAMPLE SIZE NEEDED TO OBTAIN A
[2]   ⍝   BAYESIAN INTERVAL WITH 95 PERCENT CONFIDENCE LEVEL, BASED ON
[3]   ⍝   A PRIOR UNIFORM [A,B] DISTRIBUTION.  IT ASKS THE USER TO
[4]   ⍝   INPUT THE PRIOR BOUNDS A AND B AND THE DESIRED INTERVAL SIZE.
[5]   ⍝ IT NEEDS THE APL PROGRAMS BQUAN, NQUAN AND BETA TO BE STORED.
[6]   ⍝ IT TERMINATES ITS EXECUTION WHEN THE SAMPLE SIZE IS ≥ 150.  FOR
[7]   ⍝   BIGGER NUMBERS, THE VALUE OF N IN LINE 22 MUST BE INCREASED.
[8]   ⍝ IF CONFIDENCE LEVEL DIFFERENT THAN 95 PERCENT IS REQUIRED, LINES
[9]   ⍝   28 AND 29 MUST BE CHANGED ACCORDINGLY.
[10]
[11]
[12]   'ENTER A, LOWER BOUND OF UNIFORM PRIOR'
[13]   LF←⎕
[14]   ' '
[15]   'ENTER B, UPPER BOUND'
[16]   RT←⎕
[17]   ' '
[18]   'ENTER 95 PERCENT C.L. INTERVAL SIZE (MUST BE LESS THAN B-A)'
[19]   INT←⎕
[20]
[21]   N←9
[22]  CONT:→FIN×⍳N=151
[23]   N←N+1
[24]   C←1+N×0.5×(LF+RT)
[25]   D←N+2-C
[26]   E←C,D
[27]
[28]   LO←(0.025×(E BETA RT))+0.975×(E BETA LF)
[29]   UP←(0.975×(E BETA RT))+0.025×(E BETA LF)
[30]   LO←E BQUAN LO
[31]   UP←E BQUAN UP
[32]
[33]  LOOP:→END×⍳((UP-LO)≤INT)
[34]   →CONT
[35]  END:' '
[36]   ' '
[37]   '95 PERCENT C.L. UPPER BOUND :  ',⍕UP
[38]   '       >>        LOWER       :  ',⍕LO
[39]   ' '
[40]   '       >>      INTERVAL SIZE:  ',⍕(UP-LO)
[41]   ' '
[42]   'REQUIRED SAMPLE SIZE  :  ',⍕N
[43]   →0
[44]  FIN:'SAMPLE SIZE IS GREATER THAN 150 AND EXECUTION TERMINATED'
[45]   ' '
      ∇
```

# APPENDIX B. THE APL PROGRAMS USED TO COMPUTE THE INVERSE CDF OF A BETA DISTRIBUTED RANDOM VARIABLE.

```
      ∇ BQUAN [⎕]
      ∇ V←A BQUAN P;E;U;S;D;L;Z;DENS;I;PP;M;X;F;C2;C3;C4
[1]   ⍝ IMPLEMENTATION OF CARTER, 1947, BIOMETRIKA FOR APPROXIMATE INVERSE BETA
[2]   ⍝ 11/5/86 BEST FOR A[1]≤2×A[2], AND SEEMS TO WORK FINE
[3]   ⍝ 12/27/86 ADDED 2 NEWTON-RAPHSON ITERATIONS; ADD MORE FOR GREATER ACC.
[4]    →((⌊/A)<1)/SMALL
[5]    E←NQUAN 1-P
[6]    U←⌽¯1+2×A
[7]    S←+/÷U
[8]    D←-/÷U
[9]    L←(¯3+E*2)÷6
[10]   Z←((S÷2)×E×(L+2÷S)*0.5)-D×(L+(5÷6)-S÷3)-(D*2)×((2÷S)*0.5)×E×(11+E*2)÷144
[11]   V←÷1+(÷/⌽A)×*2×Z×I←1
[12] LOOP:DENS←A[1]×(A[1]!¯1++/A)×(V*A[1]-1)×(1-V)*A[2]-1
[13]   V←V-((A BETA V)-P)÷DENS
[14]   →((I←I+1)≤2)/LOOP
[15]   →0
[16] ⍝ MY VERSION FOR THE BETA QUANTILES WHEN A⌊B<1. 12/31/86
[17] ⍝ MODIFIED 1/1/87 WITH A CORNISH-FISHER TYPE EXPANSION.
[18] ⍝ MODIFIED 1/3/87 TO USE MEAN AND STANDARD DEVIATION, AND NORMAL QUANTILE
[19] ⍝ WHEN ONE PARAMETER IS GREATER THAN ONE (FOR ONE SIDE). OTHER SIDE (OR
[20] ⍝ BOTH) USES THE DENSITY WHICH IS UNBOUNDED, FOLLOWED BY CORNISH-FISHER.
[21] SMALL:V←X←(ρ,P)ρ0
[22]   PP←P≤M←A[2]÷+/A
[23]   X[PP/⍳ρX]←((PP/P)÷(((A[1]<1),A[1]≥1)/1,A[1])×A[1]!¯1++/A)*÷A[1]
[24]   X[(~PP)/⍳ρX]←1-((1-(~PP)/P)÷(((A[2]<1),A[2]≥1)/1,A[2])×A[2]!¯1++/A)*÷A[2]
[25]   X[(X=1)/⍳ρX]←1-1E¯15
[26]   →((⌈/A)≥1)/ONE
[27] START:F←(A[1]!¯1++/A)×A[1]×(X*A[1]-1)×(1-X)*A[2]-1
[28]   C2←((1-A[1])÷X)+(A[2]-1)÷1-X
[29]   C3←(2×C2*2)+((A[1]-1)÷X*2)+(A[2]-1)÷(1-X)*2
[30]   C4←(6×C2*3)+(7×C2×(C3-2×C2*2))+((1-A[1])÷X*3)+(A[2]-1)÷(1-X)*3
[31]   F←(P-(A BETA X))÷F
[32]   V←X+F+((C2×F*2)÷2)+(C3×(F*3)÷6)+C4×(F*4)÷24
[33]   V[(V>1)/⍳ρV]←1
[34]   →0
[35] ONE:M←1-M
[36]   S←(M×(1-M)÷1++/A)*0.5
[37]   →((A⍳⌈/A)=2)/4+⎕LC
[38]   X[PP/⍳ρX]←M+S×NQUAN PP/P
[39]   X[(X≤0)/⍳ρX]←1E¯15
[40]   →START
[41]   X[(~PP)/⍳ρX]←M+S×NQUAN(~PP)/P
[42]   X[(X≥1)/⍳ρX]←1-1E¯15
[43]   →START
      ∇
```

```
      ∇ NQUAN [⎕]
      ∇ Z←NQUAN P;A;B;C;D;Q;T;S;R;F
[1]   ⍝ IMPLEMENTS ALGORITHM AS 111 BY BEASLEY   SPRINGER, APPLIED STAT, 1977
[2]   ⍝ FOR A VECTOR INPUT OF FRACTIONS, RETURNS CORRESPONDING NORMAL QUANTILES
[3]   ⍝ WITH CLAIMED ACCURACY BETTER THAN 1.5×10*¯8. FOR GREATER ACCURACY,
[4]   ⍝ ESPECIALLY FOR EXTREME P VALUES, ADD ONE OR MORE NEWTON-RAPHSON LOOPS.
[5]    →(∨/(P≤0),(P≥1))/ERR
[6]    →(∨/((|Q←,P-0.5)≤0.42))/3+⎕LC
[7]    S←Z←,Q
[8]    →EXT
[9]    T←(0.42≥|Q)/Z←,Q
[10]   →(F←((ρ,T)=ρ,P))/2+⎕LC
[11]   S←(0.42<|Q)/Q
[12]   A← 2.50662823884 ¯18.6150006252 41.39119773534 ¯25.44106049637
[13]   B← ¯8.4735109309 23.08336743743 ¯21.06224101826 3.13082909833
[14]   T←T×(((T*2)∘.*0,ι3)+.×A)÷1+((T*2)∘.*ι4)+.×B
[15]   Z[(0.42≥|Q)/ιρ,Q]←T
[16]   →(F=1)/0
[17] EXT:C← ¯2.78718931138 ¯2.29796479134 4.85014127135 2.32121276858
[18]   D← 3.54388924762 1.63706781897
[19]   S←(×S)×((R∘.*0,ι3)+.×C)÷1+(((R←(|⍟0.5-|S)*0.5)∘.* 1 2)+.×D)
[20]   Z[(0.42<|Q)/ιρQ]←S
[21]   →0
[22] ERR:'ONE OR MORE P VALUES ARE OUT OF RANGE.'
      ∇


      ∇ BETA [⎕]
      ∇ U←A BETA X;Y;W;N;OD;EV;Z;I
[1]   ⍝ 12/27/86 EVALUATES THE BETA CDF, PARAMETERS A, AT VECTOR X USING THE
[2]   ⍝ BOUVER-BARGMAN CONTINUED FRACTION AT DEPTH VARYING FROM 7 TO 21.
[3]   ⍝ 11TH ANNUAL SYMPOSIUM ON THE INTERFACE OF COMPUTER SCIENCE AND
[4]   ⍝ STATISTICS, 1978, P 325. BECAUSE OF THE RANGE OF !, +/A≤255. SEEMS TO
[5]   ⍝ GIVE A GOOD 8 OR MORE DECIMALS.
[6]    Y←X≤(A[1]÷+/A)
[7]    U←(ρ,X)ρ0
[8]    N←7++/(⌈/A)>(2×ι4),10×ι10
[9]    →((+/Y)=0)/FLIP
[10]   W←Y/X←,X
[11]   OD←W∘.×((ιN)×A[2]-ιN)÷×/(N,2)ρA[1]+ι2×I←N
[12]   EV←-W∘.×(×⌿((2,N)ρ(A[1]+0,ιN-1),(+/A)+0,ιN-1))÷×/(N,2)ρA[1]+0,ι([illegible]N-Z←1)
[13] L:Z←1+EV[;I]÷1+OD[;I]÷Z
[14]   →((I←I-1)>0)/L
[15]   U[Y/ιρU]←(÷Z)×(A[1]!¯1++/A)×(W*A[1])×(1-W)*A[2]
[16]   →((+/Y)=ρX)/0
[17] FLIP:A←⌽A
[18]   W←1-(~Y)/X
[19]   OD←W∘.×((ιN)×A[2]-ιN)÷×/(N,2)ρA[1]+ι2×I←N
[20]   EV←-W∘.×(×⌿((2,N)ρ(A[1]+0,ιN-1),(+/A)+0,ιN-1))÷×/(N,2)ρA[1]+0,ι(2×N-Z←1)
[21] L1:Z←1+EV[;I]÷1+OD[;I]÷Z
[22]   →((I←I-1)>0)/L1
[23]   U[(~Y)/ιρU]←1-(÷Z)×(A[1]!¯1++/A)×(W*A[1])×(1-W)*A[2]
      ∇
```

## LIST OF REFERENCES


1. Fishburn, Peter C., Decision and Value Theory, John Wiley & Sons, Inc., New York, 1964.

2. Lansons, Harold J., Introduction to Probability Theory and Statistical Inference, John Wiley & Sons, Inc., New York, 1982.

3. Spiegel, Murray R., Schaum's Outline of Theory and Problems of Statistics, Schaum's Outline Series, McGraw-Hill, Inc., New York, 1961.

4. Duncan, Acheson J., Quality Control and Industrial Statistics, Richard D. Irwan, Inc., Homewood, Illinois, 1986.

5. DeGroot, Morris H., Probability and Statistics, Second Edition, Addison-Wesley Publishing Company, Inc., USA, 1986.

6. Manion, Rober B., Number of Samples Needed to Obtain Desired Bayesian Confidence Intervals for a Proportion, Master's Thesis, Naval Postgraduate School, Monterey, California, March 1988.

7. Miller, I., and E. J. Freund, Probability and Statistics for Engineers, third edition, Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1972.

## INITIAL DISTRIBUTION LIST

| | | No. Copies |
|---|---|---|
| 1. | Defense Technical Information Center<br>Cameron Station<br>Alexandria, Virginia 22304-6145 | 2 |
| 2. | Library, Code 0142<br>Naval Postgraduate School<br>Monterey, California 93943-5100 | 2 |
| 3. | Professor G. F. Lindsay, Code 55Ls<br>Department of Operations Research<br>Naval Postgraduate School<br>Monterey, California 93943-5000 | 2 |
| 4. | Professor L. Johnson, Code 55Jo<br>Department of Operations Research<br>Naval Postgraduate School<br>Monterey, California 93943-5000 | 2 |
| 5. | Hellenic Navy General Staff<br>Branch B2<br>Stratopedo Papagou, Athens<br>Greece | 5 |
| 6. | LCDR Theodore C. Floropoulos<br>Agios Barbaras 7<br>176 76 Kallithea, Athens<br>Greece | 4 |
| 7. | LT Jeffrey L. Turner<br>Code 33<br>Naval Postgraduate School<br>Monterey, California 93943-5000 | 1 |